PANORAMIC VIDEO SYSTEM WITH REAL-TIME DISTORTION-FREE IMAGING

ORIGIN OF THE INVENTION

The invention described herein was made in the performance of work under a USMC contract No. M67854-03-C-1006, and is subject to the provisions of public law 96-517 (35 USC 202) in which the contractor has elected to retain title.

FIELD OF THE INVENTION

The present invention relates generally to the field of special video camera systems and more specifically to a real-time 360° panoramic video system which utilizes a panoramic annular mirror, video camera and unique unwrapping software which provides a seamless, distortion-free horizontal view of the panoramic image.

CROSS-REFERENCE TO RELATED APPLICATIONS

The present invention takes priority from provisional application Serial No. 60/485,336 filed on July 3, 2003.

BACKGROUND OF THE INVENTION

Panoramic optical systems which can be employed to provide a 360° field of view, are known. By way of example, U.S. Patent No. 6,459,451 discloses a catadioptric lens which provides a 360° field of view. Such optical systems can be used advantageously with a camera to provide a system capable of imaging an entire 360° field such as an entire room or landscape from a unitary location using a single camera without requiring scanning or stitching multiple images. However, such lenses provide an image which is not readily interpretable by observers unless the image is first "unwrapped". The image of a 360° field of view lens is annular or doughnut-shaped and is therefore distorted and not readily discernible by a human observer. It is therefore necessary to convert that image or "unwrap" it into a relatively two-dimensional format such as a horizontal view on a relatively flat medium such as physically on film or electronically on a computer screen. The unwrapping process consists of a mathematical transformation such as by conversion of each picture element or pixel and is preferably accomplished in a manner which results in little or no distortion which would otherwise reduce the quality of the resulting flat image. Such pixel-by-pixel transformations are typically very complex and require complicated and time consuming computer programs, especially for reasonable levels of resolution and images having large numbers of pixels. Consequently, it has not been possible heretofore to exploit panoramic lens technology to provide a real-time unwrapped video image with acceptable resolution.

A system which could provide real-time unwrapped video images derived from a panoramic lens and video camera would be highly advantageous for a variety of useful applications. By way of example, such a system could provide security surveillance over a continuous all-around field of view using a unitary display with just one observer. Such a system could also be mounted on a transport mechanism and used for military or police reconnaissance purposes or for robotic imaging. It could also be used for medical visualization and for traffic

2

1   awareness systems. It can be tailored to be compatible with internet

2   transmission, wireless systems and can be designed for video image

3   compression to reduce transmission bandwidth requirements. Once it becomes

4   convenient to "unwrap" a panoramic video image in real time with little or no

5   distortion and with an acceptable level of resolution, a host of useful and

6   advantageous applications become feasible and readily available.

SUMMARY OF THE INVENTION

1
2
3   The present invention in its preferred embodiment combines a panoramic
4   annular lens system (PAL), a unitary video camera and a PC-based software
5   system that unwraps a 360° video image into a seamless, distortion free
6   horizontal image image in real time. The PAL system comprises two mirrors,
7   namely, a hyperboloidal mirror and an ellipsoidal mirror interconnected by a 360°
8   circular refracting front or entrance aperture lens and having a rear or exit
9   aperture adjacent a collector lens. The PAL system of the preferred embodiment
10 has a 360° horizontal field of view and a 90° vertical field of view in a 40mm
11 diameter compact package. The invention is not limited to any particular type of
12 lens system. In fact, there are numerous lens systems for providing a 360°
13 panoramic view. The video camera may be a CCD or CMOS based device
14 having a pixel resolution of either 1280 x 1024 (high resolution) or 720 x 480
15 (NTSC). The unwrapping system is a radiometric ray tracing program carried out
16 using a computer's graphics card capabilities to produce highly efficient regional
17 transformation while minimizing software overhead. The result is real time, high
18 resolution 30 fps conversion from a spherical distorted image to a flat panoramic
19 image in Cartesian coordinates. A graphic user interface (GUI) permits selection
20 of any breaking point (any center line of the panoramic image) as well as zoom in
21 and zoom out capability and built-in calibration.

BRIEF DESCRIPTION OF THE DRAWINGS

The various embodiments, features and advances of the present invention will be understood more completely hereinafter as a result of a detailed description thereof in which reference will be made to the following drawings:

FIG. 1 is a schematic diagram of a panoramic annular lens structure which may be employed in the present invention;

FIG. 2, comprising FIGs. 2a and 2b, shows photographic top and side views, respectively of the lens structure of FIG. 1;

FIG. 3 is a photographic view of the lens structure of FIG. 1 shown integrated with a CCD camera;

FIG. 4, comprising FIGs. 4a and 4b, is a photographic illustration of a PAL image before and after clipping, respectively;

FIG. 5 is a photographic representation of the unwrapped version of the PAL image of FIGs. 4a and 4b;

FIG. 6, comprising FIGs. 6a, 6b, 6c and 6d, provides photographic views of the image of FIGs. 4a and 4b in wrapped form in two different camera resolutions and in unwrapped form in the same two resolutions, respectively;

FIG. 7 is a schematic diagram of an alternative embodiment of a catadioptric omnidirectional ultra-wide-angle camera;

FIG. 8 is a simplified illustration of panoramic stereo imaging using a double parabolic mirror;

1   FIG. 9 is a schematic diagram showing the design of a multilevel parabolic

2   mirror and camera;

3

4   FIG. 10 is a schematic diagram of panoramic imaging using a convex

5   reflecting mirror;

6

7   FIG. 11 is a schematic diagram of panoramic camera system useful for

8   day and night operation;

9

10   FIG. 12 is a schematic diagram of an annular flat mirror used in the

11   system of FIG. 11;

12

13   FIG. 13 is a schematic diagram of panoramic imager having a second

14   medium wavelength infrared optical channel;

15

16   FIG. 14 is a schematic diagram of the hemisphereic view circular

17   projection of a circular fisheye lens;

18

19   FIG. 15 is a geometric diagram of a spherical coordinate mapping of the

20   circular fisheye lens;

21

22   FIG. 16, comprising FIGs. 16a and 16b, is a geometrical representation of

23   spherical and angular mapping, respectively;

24

25   FIG. 17, comprising FIGs. 17a and 17b, is a photographic view of the

26   original and converted images, respectively, of a circular fisheye lens;

27

28   FIG. 18 is a computer screen representation of the graphic user interface

29   for real-time conversion ("unwrapping") software of the preferred embodiment of

30   the invention;

31

1   FIG. 19 is a computer screen representation of an image captured by a
2   fisheye video camera system in full frame at 30 fps;
3
4   FIG. 20 is a 1500x1000 Cartesian computer screen image converted at 30
5   fps from the fisheye image of FIG. 19 using the "unwrapping" software of the
6   present invention;
7
8   FIG. 21 is a computer screen representation similar to that of FIG. 19 but
9   showing the effect of various calibration methods for reducing distortion;
10
11   FIG. 22 is a graphical representation of radiometric ray tracing from an
12   input plane to an output plane for a pixel block;
13
14   FIG. 23, comprising FIGs 23a, 23b and 23c, is a graphical representation
15   showing transformation process from an annular image to an undistorted
16   rectangular image;
17
18   FIG. 24 is a block diagram of software flow of real-time unwrapping used
19   in the present invention;
20
21   FIGs. 25 and 26 are schematic diagrams used to explain the use of
22   panoramic imagers as rangefinders; and
23
24   FIGs. 27 and 28 are schematic diagrams used to explain the addition of
25   zoom function to a panoramic imager.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Panoramic Annular Lens (PAL)

The PAL lens is based on both reflection and refraction of light and offers panoramic 360° field of view in an ultra compact packaging of only 40 mm diameter. The PAL lens provides a vertical field of view such as -40° to +50°. As shown in FIG. 1, the panoramic lens is a piece of glass that consists of a 360° circular aperture (R1), a rear aperture (R2) connecting to a collector lens, a top mirror (H) and a circular mirror (E). The viewpoint of the "virtual camera" is at the plane (O) of the ellipsoidal mirror (E). With this geometry, the PAL sensor can view the entire 360° scene around its vertical axis BC. The vertical field of view is determined by the effective sizes and the locations of the circular mirror E and the top mirror H. Usually the viewing angle is 90° vertically.

The PAL is shown in FIGs. 2a and 2b. To maintain wide camera angle options, the PAL mounting is terminated with a C-type mount that fits most 1/3 in. and 1/2 in. pick-up devices, including CMOS and CCDs. Selection of a pick-up device is important because it defines the final image quality. The most important characteristic is resolution, which should be on the order of 1000 pixels per video line. Progressive mode pick-up devices are preferred because they eliminate temporal video field disparity. A wide range of such devices is available on the market, with the prices dropping as more market share goes to HDTV camcorders such as the JVC GR-HD1, which can record video in 720p HD format (1024x720 pixels in progressive scans at 30 fps). The PAL integrated with a camera is shown in FIG. 3.

1   The image produced by the PAL lens is circularly symmetric, as seen in

2 FIG. 4a. The PAL lens maps all the space within a 360° azimuth and 90°

3 elevation into an annular ring image. The image can still be recognized, and it

4 has relatively low geometric distortion compared to other panoramic visualization

5 systems such as hyperbolic mirror-based 360° systems. The major advantage of

6 the PAL is that it keeps vertical lines straight, significantly reducing the

7 computational complexity of image unwrapping. Only part of the image in FIG.

8 4a can usefully be unwrapped, as seen in FIG. 4b. The center and outer edges

9 of 4a do not carry any useful visual information, and are discarded. We maintain

10 maximum resolution covering the entire annular ring image by optimizing the

11 relay lens between the PAL and the camera. The unwrapped image unfolded to

12 panoramic Cartesian coordinates is shown in FIG. 5.

13

14   The PAL assembly was extensively tested on several cameras with a

15 variety of pixel resolutions. We started with a standard NTSC camera, which

16 thus had 420×240 pixels per field. Next we tested the lens with progressive scan

17 cameras, one at 480×480 pixels and one at 1024×1024. For each PAL-camera

18 combination we unwrapped the image and normalized it to the reference camera

19 resolution by bicubic scaling. FIG. 6 compares the images. FIG. 6a shows the

20 image captured by the 1024×1024 pixel camera. FIG. 6b shows a 480×480 pixel

21 image. The corresponding unwrapped images are shown in FIG. 6c for

22 1024×1024 pixels and in FIG. 6d for 480×480. A detail of both images (a picture

23 on the wall) is magnified in both cases. As expected, close examination of the

24 detail from both images shows smoother edges and better color rendering for the

25 high-resolution image.

26

27 Catadioptric Ultra-Wide-Angle Camera with Parabolic Mirror

28

29   In an alternative panoramic vision approach, a catadioptric system creates

30 omnidirectional vision by means of a parabolic mirror. The catadioptric

31 omnidirectional ultra-wide-angle camera (CUWAC) consists of a miniature digital

1   videocamera mounted in a frame and aimed directly at the apex of a parabolic

2   mirror enclosed within a transparent hemisphere, as illustrated in FIG. 7. The

3   dioptric camera lens in FIG. 7 images the focus of the parabola onto a CCD

4   imager. This camera views in all directions within a hemisphere via a parabolic

5   mirror. A CCD camera with a dioptric imaging lens faces the mirror a few inches

6   away and produces a circular image of the reflection. This resulting circular

7   image can then be converted into a normal view in any direction. However, the

8   image quality varies a great deal across the field-of-view; the system

9   magnification is greater at the center of the image and gradually decreases as

10   the image height Y increases. This causes severe image degradation at the

11   field-of-view margins. To overcome this drawback, we extended the panoramic

12   imaging concept to two concentric parabolic mirrors that differ in curvature. In

13   the two-mirror system illustrated in FIG. 8, the axes of the mirrors are collinear,

14   and coincide with the optical axis of the dioptric camera. Each of the mirrors has

15   a profile radially symmetric around this axis. The major parabolic mirror causes

16   less demagnification, and captures those parts of the hemisphere at larger

17   angles with better resolution. The minor mirror has a higher curvature and

18   higher magnification, and captures the central parts of the scene, i.e., close to

19   the optical axis.

20

21   The CUWAC parabolic optics ensure that it has a single effective center of

22   projection, a single point through which all rays from a scene must pass on their

23   way to the camera lens. That design mimics a camera that takes in only linear

24   perspective, and allows the CUWAC computer software to generate linear

25   perspective images that are free of distortion.

26

27   Two cameras with fisheye lenses or parabolic mirrors mounted back-to-

28   back can produce views of 360°, a complete sphere, for surveillance or security

29   operations. In teleconferences, such a panoramic camera can show

30   simultaneously every participant seated around a table, in either hemispheric or

31   linear perspective. It will allow a mobile robot to view hemispheric scenes.

10

1   Placed atop a concert stage or above midfield during a sports event, the ultra-

2   wide-angle camera could provide a 360° view — an entire sphere — to viewers.

3   With a joystick or mouse, the viewers could bring any view to their screens, and

4   see not only hemispheric perspective but normal, undistorted, linear perspective.

5

6   The basic design of this multilevel parabolic mirror is shown in FIG. 9.

7   The projection center C of the hyperbolic mirror coincides with focal point F. The

8   perspective camera is modeled by an internal camera calibration matrix K, which

9   relates 3D coordinates $X = [x,y,z]^T$ to retinal coordinates $q = [q_u,q_v,1]^T$

10

11
$$q = \frac{1}{z}KX \ .$$
(1)

12

13   Analysis of Panoramic Imaging System

14

15   Research in remotely operated and autonomous systems has shown the

16   usefulness of imaging that can span a very wide field-of-view. If instead of a

17   small conic view, a camera can capture almost an entire hemisphere of visual

18   information ("view-sphere") at a time, the imaging system gains several

19   advantages. First, it is not necessary to move the camera to fixate on an object

20   of interest, or to perform exploratory camera movements. Second, processing

21   global images of an environment or target scene is less likely to be affected by

22   regions of the image that contain poor information. Third, a wide field-of-view

23   eases the search for reference objects, as they do not disappear from the field-

24   of-view; it helps stabilize image processing algorithms tracking such features.

25   Fourth, a wide field-of-view makes it easier to distinguish image artifacts due to

26   rotation of the camera from image artifacts due to object translation. The ability

27   to view and image an environment panoramically is useful in applications ranging

28   over machine vision, surveillance, collision avoidance, computation of ego-

29   motion, simple and easy detection of objects moving in the environment, and

30   robotics.

1

2   Reflective optics is cost effective and robust for global imaging. A camera

3   placed below a convex reflecting surface can observe large fields-of-view (see

4   FIG. 10). The mirror profile can be designed for angular gain to extend the

5   camera viewing geometry. With an appropriately shaped mirror, cameras that

6   typically have visual fields-of-view of ~30° can now image a full 360° in azimuth

7   θ, and up to ±120° in elevation Φ. FIG. 10 illustrates how a ray reflecting off the

8   reflector surface is directed into the camera viewing cone. Such an imaging

9   device has obvious advantages: first, being a passive sensor it has minimal

10   power requirements. Second, it has the potential to be extremely robust, since

11   the sensor is purely solid state and has no moving parts. Third, curved mirrors

12   can be made relatively free of the optical distortion that is typical of lenses.

13

14   Cameras with convex mirrors can acquire imagery instantly at video rates;

15   they can be compact, and can cost relatively little to produce. Commercial

16   "fisheye" lenses tend to be much more costly and bulkier than mirrors. In

17   addition, camera optics based on convex mirrors can have well defined

18   mathematical relationships that can be coded into the image processing and data

19   filtering to map the curved geometry of the view-sphere onto the 2D planar pixel

20   array. There are no simple and cost effective purely optical means for correcting

21   the image deformation that occurs in going from a 3D representation to a 2D

22   representation. A fundamental difficulty with panoramic imaging using a curved

23   reflective surface is that image resolution depends on position within the image.

24   In images from standard spherical convex reflectors, resolution depends upon

25   elevation. Visual patches at high elevations are quite different in resolution from

26   those near the horizontal, because they capture smaller solid angles of visual

27   space than do equatorial patches. Designing the mirror profiles to be

28   equiangular, transforms a curved image into a cylindrical projection, preserving a

29   linear relationship between the angle of incidence of light onto the mirror surface

30   and the angle of reflection into the camera with respect to the center of the

31   detector array. This ensures that the camera maintains uniform resolution of the

12

1 environment in the vertical plane independent of elevation angle, which is very

2 important to high quality panoramic imaging. Left unaccounted for in both mirror

3 design and image processing, vertical nonuniformity causes poor resolution

4 across a given target scene.

5

6 POC Panoramic Imager for Day/Night Operation

7

8 FIG. 11 illustrates a modular visible/infrared camera system. Light from

9 the scene is incident on a hyperbolic mirror. The surface profile of this mirror

10 (i.e., conic constant, radius of curvature, and aperture size) is designed in such a

11 way that the focus of the hyperbolic curve acts as the camera projection center,

12 where all rays appear to intersect. Alignment of the mirror with the cameras in

13 this system is critical to maintain the linear relationship between the elevation

14 and camera viewing angles. Those rays satisfying the single viewpoint

15 relationship are reflected by the hyperbolic mirror surface, and are incident on an

16 annular flat mirror (see FIG. 12) that is oriented at 45° with respect to the nadir.

17 Half of the light hitting the annular mirror passes through the clear aperture

18 (lower elevations of the viewing geometry) within the central portion of the mirror,

19 and half the light (higher elevations of the viewing geometry) is reflected at 90°.

20

21 The light propagating along each optical path is collected by a zoom lens.

22 The video zoom lens for this optical system is a commercial off-the-shelf product

23 with a focal length varying from 8 mm to 48 mm, a working distance that ranges

24 from 1.2 m to ∞, and compatibility with 1/2 in. format detectors, and has F

25 numbers that range from F1.2-16 and angular fields-of-view from 44.6° to 8°.

26 The two zoom lenses enable us to independently adjust each arm of the sensor.

27 They need not be set to the same zoom magnification (i.e., the blur spot size can

28 be set to fill up the pixel array); this can improve the resolution in each portion of

29 the visual image, which has the benefit of enabling the system to equalize

30 resolution as a function of viewing elevation for the panoramic imager. The

31 minimum blur spot size for the panoramic imager with this zoom lens is estimated

1   to be ~1.5 to 2 times the diffraction limit. The light from each zoom lens is
2   imaged onto a commercially available 3 megapixel, 1/2 in. format, CMOS silicon
3   detector chip. The number of pixels in each detector array is 2048x1520, with a
4   7 μm linear pitch. The larger pixel size improves the low light sensitivity of the
5   camera chip to ~0.05 lux with reduced fixed pattern noise. According to the
6   manufacturer's specifications, the SNR for this camera is 78 dB. The camera
7   operates in noninterlaced mode (progressive scan), and produces full frame
8   readouts at video rates of 30 frames per second. Full asynchronous image
9   capture with programmable partial scan (region-of-interest mode of operation)
10   gives these cameras the flexibility for numerous daytime/nighttime applications.
11   Both color and monochrome versions of the camera are available. In the color
12   version the overall resolution is reduced by the Bayer color filter; the resolution is
13   about 1/3 for each of the primary colors.
14
15   Because silicon is sensitive into the near infrared region (700 nm to 1100
16   nm) of the spectrum, the imager can be used for nighttime surveillance.
17   Moreover, the panoramic imager is designed to be modular so that a second
18   channel can easily be introduced to extend the nighttime vision capability into the
19   mid-wave infrared (3 to 10 μm) region. This design is envisioned with a flat
20   mirror, coated for >98% reflectivity over the MWIR spectrum. The 45° orientation
21   of the flat mirror directs the light toward a second optical channel (see FIG. 13).
22
23   The rest of the optical layout would be similar to that described previously,
24   with the addition of PtSi or HgCdTe detectors and infrared zoom lens assemblies
25   to detect the infrared scene. Note that these infrared focal plane arrays are only
26   QVGA scale (320x240 pixels), with a linear pitch of 12 μm, so overall resolution
27   would be reduced. However, a multicolor panoramic imager could track targets
28   under conditions that would be beyond the capabilities of the silicon-based
29   detectors. At night infrared irradiance is about two orders of magnitude greater
30   than that in the visible spectrum under moonlight conditions. Additionally one

1   may apply sophisticated image interpolation techniques to increase the image

2   resolution.

3

4   <u>Mathematical Foundation for Panoramic Image Unwrapping</u>

5

6   The circular fisheye lens projects a hemispheric view of the surroundings

7   into a circular image as shown in FIG. 14. The panoramic image is a 180°

8   fisheye projection. The projected panoramic image covers a full 180°

9   horizontally, but because of cropping of the frame it covers substantially less

10   vertically, ~ 135°.

11

12   In constructing the unwrapping process, the unit assigned to the limiting

13   circle of the 180° fisheye projection is radius, and its center is chosen as the

14   image origin. Points in the image are assigned polar coordinates (r,θ) and

15   converted to spherical coordinates with angular coordinates θ and ϕ, where θ is

16   longitude and ϕ is the angle from the axis of projection as in Eq. (2). FIG. 15

17   geometrically illustrates the spherical mapping of a circular fisheye image. The

18   transformation from polar to spherical coordinates keeps θ the same and

19   transforms r into ϕ. FIG. 16 shows the angles of mapping coordinates (FIG. 16a)

20   and a geometrical representation of angular coordinate conversion (FIG. 16b).

21

22
$$
\begin{pmatrix} x_b \\ y_b \\ z_b \end{pmatrix} = \begin{pmatrix} \cos\alpha & 0 \\ \sin\alpha\sin\beta & \cos\beta \\ -\sin\alpha\cos\beta & \sin\beta \end{pmatrix} \begin{pmatrix} x_d \\ y_d \end{pmatrix} + \begin{pmatrix} \sin\beta \\ -\sin\alpha\cos\beta \\ \cos\alpha\cos\beta \end{pmatrix}
\tag{2}
$$

23

24   Then we can map the hemicube to the fisheye image, and from this we

25   can convert a 180° fisheye image (see FIG. 17a) into a normal perspective

26   image, with the result shown in FIG. 17b. The mapping equations used for

27   transformations of coordinates are:

28

$$\begin{pmatrix} \phi_b \\ \varphi_b \end{pmatrix} = \begin{pmatrix} \arctan\left(\dfrac{y}{x}\right) \\ \arctan\left(\dfrac{\sqrt{x_b^2 + y_b^2}}{z}\right) \end{pmatrix} \tag{3}$$

$$\begin{pmatrix} x_v \\ y_v \end{pmatrix} = \begin{pmatrix} r_v \cos\theta \\ r_v \sin\theta \end{pmatrix} \tag{4}$$

$$\begin{pmatrix} \theta_v \\ r_v \end{pmatrix} = \begin{pmatrix} \phi_b \\ r\varphi_b \end{pmatrix} \tag{5}$$

The mapping pipelines the following steps for continuous operation of incoming images:

1. Image plane to angular coordinate
2. Angular coordinate to spherical coordinate
3. Find inverse transformation

$$(x_d, y_d) \xrightarrow{(\alpha,\beta)} (x_b, y_b, z_b) \rightarrow (\phi_b, \varphi_b) \xrightarrow{r} (\theta_v, r_v) \rightarrow (x_v, y_v) \tag{6}$$

Some of the necessary equations involve spherical coordinates. The angles θ and ϕ in the following equations are related to a canonical Cartesian (x,y,z) coordinate frame by:

$$x = r \cdot \sin(\delta) \cdot \cos(\theta) \tag{7}$$

$$y = r \cdot \sin(\delta) \cdot \sin(\theta) \tag{8}$$

$$z = r \cdot \cos(\theta) \ , \tag{9}$$

and their inverse:

$$r^2 = x^2 + y^2 + z^2 \tag{10}$$

$$\cos(\theta) = x / \sqrt{x^2 + y^2} \qquad\qquad (11)$$

$$\sin(\theta) = y / \sqrt{x^2 + y^2} \qquad\qquad (12)$$

$$\cos(\phi) = z / r \qquad\qquad (13)$$

$$\sin(\phi) = \sqrt{x^2 + y^2} / r. \qquad\qquad (14)$$

Real-time Panoramic Video Conversion Software

This section discusses the fundamentals of video mapping software architecture and design issues. The conversion system is coded and debugged based on Microsoft Windows Video technology and additional video manipulating software architecture. The performance and stability of the software have been optimized for real-time video conversion software.

Designs and Functionalities of Video Mapping Software

The preferred embodiment of the invention comprises real-time panoramic video conversion software to convert video from circular polar coordinates to Cartesian panoramic video with 2000×1000 video resolution at 30 frames per second.

The real-time panoramic conversion software has been developed in Microsoft Direct3D and DirectShow. Microsoft Direct3D has full capabilities for rendering and mapping images in real time. Direct3D can directly access and manipulate video memory without calling upon operating system services, so the graphics can be manipulated in hardware. The following lists summarize the capabilities of Direct3D.

1 Direct3D functionality

2 • Device-dependent access to 3D video-display hardware in
3   a device-independent manner

4 • Support for 3D z-buffers

5 • Switchable depth buffering

6 • Transformation and clipping

7 • Access to image-stretching hardware

8 • Exclusive hardware access

9 • Immediate access to the transformation, lighting, and
10   rasterization 3D graphics pipeline

11 • Software emulation if hardware acceleration is not available

12

13 Direct3D low level functionality

14 • 3D coordinate systems and geometry

15 • Shading techniques

16 • Matrices and transformations

17 • Vectors and vertices

18 • Copying surfaces

19 • Page flipping and back buffering

20 • Rectangles

21

22 Direct3D application level functionality

23 • Bump mapping

24 • Environment mapping

25 • Geometry blending

26 • Indexed vertex blending

27 • Patches

28 • Point sprites

29 • Procedural pixel shader

30 • Procedural vertex shaders

31 • Quaternions

1   •   Spotlights

2   •   Tweening

3   •   Vertex blending

4   •   Volume textures.

5

6 Microsoft introduced new technology to apply Direct3D to video

7 applications gluelessly for real-time manipulation of video with mapping,

8 blending, textures, and shadings. The following highlights DirectShow

9 technology.

10

11   •   Architecture for streaming media

12   •   High-quality playback of multimedia streams

13   •   File based

14   •   Network stream

15   •   Universal decoding capability

16   •   Glueless interface with other DirectX technology

17   •   Automatic detection of hardware acceleration support

18   •   Common Object Model (COM)-based interface.

19

20 The real-time video software was developed around the core functions of

21 Microsoft Direct3D and DirectShow, but the innovative and unique architectural

22 and hierarchical development of this software is the first in the multimedia world

23 that can convert and display panoramic video in real time without noticeable

24 latency.

25

26 Customizing Microsoft Video Mixing Renderer

27

28 The Video Mixing Renderer (VMR) is a new DirectShow filter, available

29 only for Windows XP Home Edition and XP Professional, replacing both Overlay

30 Mixer and Video Renderer, and adding many new mixing features. In terms of

1   both performance and breadth of features, the VMR represents the new

2   generation in video rendering on the Windows platform.

3

4   VMR supports the following new features:

5

6   •   Real mixing of multiple video streams, taking advantage

7   of the alpha-blending capabilities of Direct3D hardware

8   devices.

9   •   The ability to plug in your own compositing component

10   to implement effects and transitions between video

11   streams entering the VMR.

12   •   True windowless rendering. It is no longer necessary

13   to make the video playback window a child of the

14   application window to play video back. The VMR's new

15   windowless rendering mode makes it easy for applications

16   to host video playback within any window without having

17   to forward window messages to the renderer for renderer-

18   specific processing.

19   •   A new renderless playback mode, in which applications

20   can supply their own allocator component to get access to

21   the decoded video image prior to its being displayed on

22   the screen.

23   •   Improved support for PCs equipped with multiple monitors.

24   •   Support for Microsoft's new DirectX Video Acceleration

25   architecture.

26   •   Support for high-quality video playback concurrently in

27   multiple windows.

28   •   Support for DirectDraw Exclusive Mode.

29   •   100% backward compatibility with existing applications.

30   •   Support for frame stepping and a reliable way to capture

31   the current image being displayed.

1   •   The capability for applications to easily alpha-blend their

2   own static image data (such as channel logos or UI

3   components) with the video in a smooth flicker-free way.

4

5   The VMR depends entirely on the graphics processing capabilities of the

6 computer display card; the VMR does not blend or render any video on the host

7 processor, as doing so would greatly impact the frame rate and quality of the

8 video being displayed. The new features offered by the VMR, particularly

9 blending of multiple video streams and/or application images, depend strongly on

10 the capabilities of the graphics card. Graphics cards that perform well with the

11 VMR have the following hardware support built in:

12

13   •   Support for YUV and "non-power of 2" Direct3D texture surfaces.

14   •   The capability to StretchBlt from YUV to RGB DirectDraw surfaces.

15   •   At least 16 MB of video memory if multiple video streams are to

16   be blended. The actual amount of memory required depends on

17   the image size of the video streams and resolution of the display

18   mode.

19   •   Support for an RGB overlay or the capability to blend to a YUV

20   overlay surface.

21   •   Hardware accelerated video decoding (support for DirectX

22   Acceleration).

23   •   High pixel fill rates.

24

25   In our conversion software, we specifically customized VMR renderless

26 mode to maximize the capability and flexibility of the software to better

27 manipulate the controlling parameters. VMR renderless mode features a

28 customized allocator for the customized rendering surface, and a customized

29 renderer for the customized rendering mode.

1

2   In renderless playback mode, the application

3

4   •   Manages the playback window.

5   •   Allocates the DirectDraw object and the final frame buffer.

6   •   Notifies the rest of the playback system of the DirectDraw object

7   being used.

8   •   "Presents" the frame buffer at the correct time.

9   •   Handles all resolution modes, monitor changes, and "surface

10   losses" – advising the rest of the playback system of these events.

11

12   The VMR

13

14   •   Handles all timing related to presenting the video frame.

15   •   Supplies quality control information to the application and the

16   rest of the playback system.

17   •   Presents a consistent interface to the upstream components of

18   the playback system, which are not aware that the application is

19   performing the frame buffer allocation and the rendering.

20   •   Performs any video stream mixing that may be required prior to

21   rendering.

22

23   Basically, the conversion software calls various functions of VMR and

24 customized DirectX surfaces to make them fit our specific purpose, which is real-

25 time non-linear image transformation with streaming video in progress. We

26 purposely coded this software with VMR-7 for Windows XP only. DirectX 9 with

27 VMR-9 code migration can be made software compatible with other operating

28 systems such as Windows 9x and Windows 2K as well as Windows XP.

1 Real-time Panoramic Unwrapping Software

2

3 The real-time conversion software implements Direct3D Immediate Mode
4 with geometrical relationships to convert spherical images to Cartesian images in
5 arbitrary perspective. The world management of Immediate Mode is based on
6 vertices, polygons, and commands that control them. It allows immediate access
7 to the transformation, lighting, and rasterization 3D graphics pipeline. Image
8 conversion is applied to primitives so that there is no intervening overhead from
9 other interfaces and direct access hardware functions. We tested our software
10 with Nvidia GeForce, ATI Radeon, and Intel low-profile VGA chips. The final
11 code is compatible with most video acceleration chips and processors, so it can
12 be used with major hardware platforms. FIG. 18 shows the graphic user
13 interface (GUI) for the conversion software. FIG. 19 shows fisheye video in the
14 GUI. FIG. 20 shows the corresponding panoramic video converted from the
15 spherical fisheye image in real time (on the fly) at the full 30 frames/second. The
16 unwrapping algorithm may be optimized to alleviate distortion near frame edges.
17 FIG. 21 shows enhanced functions of the player.

18

19 The following list summarizes the player capabilities and functions:

20

21 • Universal playback capability (MPEG, AVI, and more)
22 • Bitmap capture (right-button click)
23 • Calibration by clicking three points on a circle (shows circle in red)
24 • Adjustments of aspect ratio for non-square pixel digitizers
25 • Change center of view in 360° panoramic viewing mode
26 • Zoom, pan, and tilt
27 - F1 and F2 for zoom-in and -out
28 - Arrow buttons for pan and tilt
29 • Status bar indicates movie size, capture device, playing time,
30 resolution of movie and display, performance, and viewing angles
31 • Capture (and DV) device properties.

1

2 And performance parameters are:

3

4 • Panoramic and 360° view: ~30 frames/second with anti-aliasing

5 and anisotropic filters

6 • 180° view: ~20 frames/second and up depends on processor

7 and video card with anti-aliasing and anisotropic filters

8 • Video resolution – full screen up to 2048x1536 pixels.

9

10 Radiometric Ray-Tracing (R^2T)

11

12 R^2T yields a single-valued mapping of radiometric quantities such as

13 luminance (brightness) from an input plane to an output plane as shown in FIG.

14 22 for any continuous optical medium.

15

16 This procedure is illustrated in FIGs. 23a, 23b and 23c, progressing from

17 an annular image (a) to a distorted unwrapped image (b), and then to a corrected

18 unwrapped image (c). Transformation from a distorted (b) panoramic image of a

19 French window into a perfect (c) panoramic image is by R^2T, is based on a priori

20 ray-tracing, equivalent to transformation from a curved system of coordinates into

21 the Cartesian coordinate system. To perform this operation in real time, we

22 divide the 120° horizontal viewing area into a large number of pixel blocks, 10x10

23 or more, and develop a uniform transformation for each block separately.

24

25 Hardware Acceleration

26

27 Conventional omni-view systems use an image by image conversion

28 process to generate transformed (or unwrapped) video with or without hardware

29 acceleration. Yet, conventional transformations (of commercial off the shelf

30 software) are done pixel by pixel and require tremendous CPU power to

31 transform the spherical or cylindrical images to Cartesian or plain view images.

1   The Applicant's approach to this problem is to increase the performance many
2   fold (typically about ~30 times faster than any existing software in current
3   market), and yet retain the quality of images or video. Animation applications
4   require tremendous processing power for manipulating images because they
5   must provide the real time conversions of perspective, zoom level, morphing,
6   rendering, and so on. Many competing video card manufacturers have been
7   developing these 3D animation engines for Video card (or Graphics Processing
8   Unit) with very high performance graphical performance for better games. We
9   have found that using these high-performance features of GPUs or 3D engines
10   makes it possible to achieve real-time performance on omni-view systems for
11   real-time video conversions.
12
13   The inventive software utilizes vertex based transformation rather than
14   pixel by pixel processing. Each vertex includes coordinates, color, and image
15   plane information for perspectives. Number of vertices for a single image is
16   limited to 65536 because we did not find any quality enhancement for more
17   vertices than 65536. The following Graphics Processing Units' (GPU) internal
18   functions are used for hardware acceleration: Lighting, Geometry Blending,
19   Alpha, Anisotropic filter or linear filters for distortion correction, 3D textures, Cube
20   textures, Clamping, and Vertex and pixel Pipelining.
21

22   • Lighting: provides detailed brightness and optical calibration
23   • Geometry and Vertex Blending: increases the realism of transformed
24   image
25   • Alpha Blending: provides the character and drawing overlays to primary
26   video
27   • Anisotropic filter: increases the quality by minimizing transforming
28   distortion
29   • 3D textures: easier for 3D coordinate transformation
30   • Cube textures: for perfect transformation on arbitrary perspective
31   • Clamping: for out-of-bound image control

1   • Vertex and pixel Pipelining: increases image manipulation in many orders
2   of performance.

3

4   In summary, the actual image manipulation of colors, lines, points, and
5   perspective changes are all done in hardware or in GPU more specifically.
6   Furthermore, video presentation is also done in hardware with no intervention to
7   other functionalities such as 3D manipulation of vertices. The only CPU load is
8   to calculate the vertex coordinate changes according to the governing equations,
9   e.g. transforming spherical coordinates to Cartesian coordinates.

10

11   Software Flow

12

13   As shown in FIG. 24, the software comprises four modules: 1) video
14   capturing, 2) image conversion, 3) presenting image as image manipulation
15   module, and 4) rendering image on video surface. The software provides many
16   video capturing features such as DV capture, video capture with any Microsoft
17   ActiveMovie compliances (16 bit), video capture with WDM (windows driver
18   model – 32 bit) drivers for video streaming, and third party capture drivers that
19   are recognized by Windows operating system.

20

21   Video capture module often introduces significant overheads and
22   resources for software. However, providing proper pipelining and buffering
23   avoids those problems.

24

25   The image conversion module converts incoming bitmaps in pipelined
26   fashion with multithreading (or super pipelining), so that minimum delay is
27   guaranteed during the process.   In this module, Microsoft Direct3D and DirectX
28   functions are utilized for image conversions and manipulation of video memory.
29   The image manipulation is done in primitive level rather than in application level,
30   so that we can have maximum level of programmability and flexibility. In

26

1   addition, the optimization of performance is possible but the coding is extremely

2   tedious because it requires programming in C++, C, and assembly languages.

3

4   The presentation module is responsible for preparing video, bitmap

5   capture, calibration (feed-backed to image conversion module), and drawing

6   circles or showing performance information on top of video.

7

8   Finally, the rendering module is heavily dependent on hardware (video

9   engine) and mostly done by using built-in functions of Microsoft DirectShow. This

10   module sends the final image streams to video memory.

11

12   FIG. 24 illustrates the overall software architecture for our real-time omni-

13   view video software of the presently preferred embodiment.

14

15   Panoramic Rangefinder

16

17   Mirror-based panoramic omnidirectional imagers have a blind zone

18   surrounding the vertical axis (see FIG. 25). Consequently, two such imagers can

19   be positioned along the same vertical axis, one above the other, without any

20   mutual obscuration of the fields of view. This use of two such imagers produces

21   a stereo effect and provides an ability to retrieve the distance to an object based

22   upon parallax angle measurement (see FIG. 26).

23

24   Panoramic Imager With Zoom

25

26   Mirror-based panoramic imagers form an image in two steps. In the first

27   step the omnidirectional object space is transferred into an imaginary intermedial

28   ring image above the mirror (see image A' of FIG. 27). In the second step the

29   projection lens transfers the imaginary intermedial image to the real image at the

30   receiver plane (see image A" of FIG. 27). Zone 0'0 can be used for the direct

31   view of the upper field zone with zoom ability. An additional lens (negative zoom

1   lens of FIG. 28) creates an imaginary image of the upper conical object field at

2   the plane 0'0, of FIG. 28. Then a projection lens fills zone 0"0", at the receiver

3   plane with the real image of the intermedial image.

4

5   Having thus disclosed preferred embodiments to illustrate the various

6   unique features of the invention, it will now be apparent that a panoramic video

7   system according to the present invention may be implemented in various ways,

8   some of which are not necessarily described herein. Therefore, the scope hereof

9   is not to be limited to the specific disclosed embodiments, but only by the

10   appended claims and their equivalents.

11

12   We claim:

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31